

# 1 Einleitung

## *Introduction*

Hersteller industrieller Produkte sind angesichts eines permanenten Wettbewerbsdrucks zu Effizienzsteigerungen gezwungen. Dabei rücken die Produktivität und das operative Betriebsergebnis als primäre Zielgrößen in den Hintergrund, während Nachhaltigkeit in den Bereichen Umwelt, Soziales und Unternehmensführung fokussiert wird [BERG21]. Gleichzeitig ist eine Steigerung der Resilienz für produzierende Unternehmen unabdingbar, da wirtschaftliche Rezessionen, ausgelöst durch z. B. die COVID-19-Pandemie, aufgrund der engen Vernetzung nationaler Wirtschaftssysteme in kürzester Zeit, zumeist unerwartet und global eintreten können [ROME18, S. 2-3].

Im Zuge der fortschreitenden Digitalisierung im produktionstechnischen Umfeld sind immer mehr zeitbasierte Live-Daten des Fertigungsprozesses bereits verfügbar oder zumindest ohne hohe Investitionskosten, bspw. über die numerische Steuerung (engl.: Numerical Control; NC), abgreifbar. Für deren industrielle Verwertung werden zunehmend auch die infrastrukturellen Voraussetzungen bspw. im Rahmen des Exzellenzclusters *Internet of Production* geschaffen, um eine durchgehende Datenverfügbarkeit sicherzustellen [SCHU20]. Die Analyse der Daten bietet vielversprechende Möglichkeiten, die Grenzen konservativer Optimierungsansätze zu überwinden und somit den genannten Aufgaben zu begegnen, da die Daten stochastische und externe Effekte abbilden. Der hohe Abstraktionsgrad der Daten stellt hierbei eine Herausforderung dar. Einzelne Signale, wie z. B. Motorströme und Achsbeschleunigungen, sind für den Menschen in ihrem Zusammenwirken nur selektiv zu erfassen. Die Prozessdaten werden daher vermehrt modellbasiert betrachtet, also domänenübergreifend aggregiert und durch physikalische Modelle veredelt, um höherwertige Informationen, wie z. B. Prozesskräfte oder statische Abdrängungen, abzuleiten. Im Anschluss werden diese Informationen zur Analyse durch einen menschlichen Experten visualisiert oder im Sinne einer Regelung in den Prozess zurückgeleitet. [BREC17b]

Allerdings werden prozessbezogene Entscheidungszeiträume immer kürzer und Ursache-Wirkungs-Beziehungen immer komplexer. Daher steigt die Nachfrage nach datengetriebenen Lösungen zur eigenständigen Entscheidungsfindung, wie bspw. Methoden der Künstlichen Intelligenz (KI) [BREC20a]. Die große Anzahl solcher Ansätze in der Produktionstechnik verdeutlicht jedoch, dass sie auf spezifische Probleme, wie z. B. die Vorhersage der Werkzeuglebensdauer [KARA16], zugeschnitten sind [MOEH20]. Die Komplexität der Produktionsumgebung, das Ausmaß veränderlicher Randbedingungen sowie die Heterogenität von Produkten und Werkzeugmaschinen erfordern gleichzeitig eine Vielzahl maßgeschneiderter Eingangsdaten, die nur in der Großserienfertigung unter immer gleichen Voraussetzungen zur Verfügung stehen.

Daraus ergibt sich der Bedarf, das Anwendungsfeld datengetriebener Lösungen zur eigenständigen Entscheidungsfindung über unterschiedliche Prozesse hinweg zu erweitern und deren Übertragbarkeit im Hinblick auf kleinere Losgrößen zu erhöhen, der in folgender Fragestellung resultiert:

Wie können während des Fertigungsprozesses aufgezeichnete Zeitreihendaten automatisiert vorverarbeitet und strukturiert werden, sodass sie zu einem späteren Zeitpunkt zielgenau gefunden und unmittelbar extrahiert werden können, um zur Entscheidungsfindung durch datengetriebene Ansätze im spezifischen Anwendungskontext beizutragen?

Vor dem Hintergrund dieser Fragestellung wird in dieser Arbeit ein Lösungsweg zur datengetriebenen Strukturierung von Zeitreihendaten aus NC-Zerspanprozessen mit Fokus auf Fräs- und Bohrbearbeitungen erarbeitet. Hierzu wird zunächst der Stand der Technik in Forschung und Industrie untersucht (Kapitel 2). Im Zuge dessen wird herausgestellt, dass eine Vielzahl leistungsfähiger Methoden aus dem Bereich des Data-Mining existieren, die in der Produktionstechnik bislang nur individuell zur Lösung spezifischer Herausforderungen herangezogen werden, obwohl die übergreifende Wissensextraktion aus den anfallenden Daten erst durch die Verknüpfung unterschiedlicher Verfahren realisiert werden kann. Daher werden Data-Mining-Algorithmen entwickelt und kombiniert, um die lokale Differenzierung der Fertigungsprozesse zu ermöglichen. Anhand der so entstehenden Teilprozessabschnitte können zeitbasierte Daten durch maschinelle Lernverfahren kontextadaptiv aggregiert werden. Darauf aufbauend wird ein hybrides Datenanalysewerkzeug [VAN21] geschaffen, mit dem der Mensch sowohl auf Basis der durch die Algorithmen identifizierten Informationen als auch mithilfe seines Expertenwissens nach spezifischen Daten suchen und diese flexibel kombinieren kann. So wird der menschliche Experte in die Lage versetzt, einerseits maßgeschneiderte Daten über mehrere Prozesse hinweg für weiterführende Modelle zusammenzustellen und andererseits einen einzelnen Prozess z. B. hinsichtlich seines Produktivitätsoptimierungspotenzials zielgerichteter analysieren zu können (Kapitel 3 bis 7). Die zugrundeliegende Methodik ist grundsätzlich datenunabhängig anwendbar und daher auf unterschiedliche NC-Prozesse übertragbar, wird aber im Rahmen dieser Arbeit anhand von Fräsprozessen entwickelt und illustriert. Das Gesamtkonzept wird in Form eines Softwaredemonstrators mitsamt der notwendigen Hardwareinfrastruktur prototypisch realisiert, bevor der Lösungsansatz anhand realer NC-Zerspanprozesse experimentell validiert und hinsichtlich seiner Wirksamkeit evaluiert wird (Kapitel 8). Den Abschluss dieser Dissertationsschrift bilden eine Zusammenfassung und ein Ausblick auf zukünftige Forschungsaktivitäten (Kapitel 9).

Die vorliegende Arbeit entstand im Rahmen des Forschungsprojekts *Semantische Datenanalyse für Zerspanprozesse (SeDaZ)* (Förderkennzeichen 22277 N), das über die Arbeitsgemeinschaft industrieller Forschungsvereinigungen (AiF) im Rahmen des Programms zur Förderung der industriellen Gemeinschaftsforschung vom Bundesministerium für Wirtschaft und Klimaschutz (BMBK) aufgrund eines Beschlusses des Deutschen Bundestages gefördert wird, und des Exzellenzclusters *Internet of Production* (Förderkennzeichen 390621612), das durch die Deutsche Forschungsgemeinschaft (DFG) im Rahmen der Exzellenzstrategie des Bundes und der Länder gefördert wird.

## Introduction

Manufacturers of industrial products are forced to increase their efficiency in the face of permanent competitive pressure. In this context, the productivity and the operating profit as primary targets move into the background, while sustainability regarding environmental, social and governance criteria is focused [BERG21]. At the same time, increasing resilience is indispensable for manufacturing companies, as economic recessions, triggered for example by the COVID-19 pandemic, can occur in a very short time, mostly unexpectedly and globally due to the close interconnectedness of national economic systems [ROME18, S. 2-3].

In the course of advancing digitalization in production environments, more and more time series live data of the manufacturing process are already available or can at least be recorded without high investment costs, for instance via the numerical control (NC). The infrastructural prerequisites for their industrial exploitation are also being created increasingly, for example within the framework of the cluster of excellence *Internet of Production*, in order to ensure continuous data availability [SCHU20]. The analysis of the data offers promising possibilities to overcome the limitations of conservative optimization approaches and thus to address the aforementioned tasks, since the data incorporates stochastic and external effects. The high level of abstraction of the data poses a challenge here. Humans can only selectively understand the interactions of individual signals, such as drive currents and axis accelerations. Process data are, therefore, increasingly examined in a model-based manner, i.e. aggregated across different domains and enriched by physical models, in order to derive information of higher value, such as process forces or static displacements. Subsequently, this information is visualized for analysis by a human expert or fed back into the process in terms of a control loop. [BREC17b]

However, process-related decision periods become progressively shorter and cause-effect relationships increasingly complex. Therefore, a growing demand for data-driven solutions for autonomous decision making, for instance methods based on Artificial Intelligence, exists [BREC20a]. The large number of such approaches in production engineering illustrates, though, that they are tailored to specific problems, for example tool life prediction [KARA16; MOEH20]. The complexity of production environments, the extent of variable boundary conditions as well as the heterogeneity of products and machine tools simultaneously require an ample amount of customized input data, which is only available in large-scale production under immutable conditions.

This results in the need to expand the scope of data-driven solutions for autonomous decision making across different processes and to increase their transferability with regard to smaller batch sizes, which results in the following question:

How can time series data recorded during the manufacturing process be pre-processed and structured in an automated way, so that they can be found accurately at a later stage and extracted immediately, in order to contribute to decision-making through data-driven approaches in a specific application context?

Against the background of this question, an approach for the data-driven structuring of time series data from NC machining processes with a focus on milling and drilling operations is developed in this thesis. To this end, the state of the art in research and industry is examined first (Chapter 2). Hereby, it is highlighted that a large number of powerful methods from the field of data mining exist, which have so far only been used individually in production engineering to solve specific challenges, although the overarching knowledge extraction from the accumulating data can only be realized by linking different methods. Therefore, data mining algorithms are developed and combined to enable the local differentiation of manufacturing processes. With the resulting sub-process segments, time series data can be aggregated in a context-adaptive manner using machine-learning techniques. On this basis, a hybrid data analysis tool [VAN21] is created, which allows humans to query and flexibly combine specific data with the help of both the information identified by the algorithms and their expert knowledge. In this way, the human expert is enabled, on the one hand, to compose customized data across several processes for further models and, on the other hand, to analyze a single process in a more targeted manner, for example regarding its productivity optimization potential (Chapters 3 to 7). The underlying methodology is basically data-independently applicable and therefore transferable to different NC processes, but is developed and illustrated in the context of this work on the basis of milling processes. The overall concept is prototypically implemented in the form of a software demonstrator including the necessary hardware infrastructure, before the approach is experimentally validated and evaluated with respect to its effectiveness based on real NC machining processes (Chapter 8). The thesis concludes with a summary and an outlook on future research activities (Chapter 9).

The present thesis originated within the context of the research project *Semantic Data Analysis for Machining Processes (SeDaZ)* (funding code 22277 N), which is funded through the German Federation of Industrial Research Associations (AiF) within the program for the promotion of joint industrial research by the German Federal Ministry for Economic Affairs and Climate Action (BMBK) based on a resolution of the German Bundestag, and the cluster of excellence *Internet of Production* (funding code 390621612), which is funded by the German Research Foundation (DFG) as part of the excellence strategy of the German federal and state governments.